TECHNICAL REPORT NO. CS 75002-R

CONVERGENCE OF COLUMN GENERATION FOR SEMI-INFINITE PROGRAMS IN THE PRESENCE OF EQUALITY CONSTRAINTS

by

HARVEY J. GREENBERG

Department of Computer Science
Virginia Polytechnic Institute and State University
Blacksburg, Virginia 24061

February 1975

ABSTRACT

A convergence theorem is presented for the standard column generation algorithm which embodies GLM. The primary extension of earlier published theorems is the allowance of equality constraints. A related stability theorem is introduced to demonstrate robustness.

Computing Reviews No. 5.41

Keywords: Optimization, Lagrange multipliers, Mathematical Programming, Resource allocation.

INTRODUCTION

We are concerned with solving semi-infinite linear programs using column generation. This embodies the Generalized Lagrange Multiplier method (GLM) when the multipliers are searched by linear programming. Our main result may be considered an extension of Nemhauser and Widhelm's (6) convergence theorem. They extended the earlier work of Brooks and Geoffrion (1) who first observed the relation to the Dantzig-Wolfe decomposition principle. Subsequent analysis by Murphy (5) showed when old columns (previously generated) may be dropped while preserving convergence. Murphy's laws remain intact under the more general theory herein described.

Earlier analysis (viz., (1) and (6)) assumed inequality constraints satisfying Slater's inferiority condition (7), so their hypothesis could not be satisfied by the usual transform of an equality constraint to a pair of inequalities. Further, their theorems are stated in such a way that the presence of gaps leaves the question of convergence (and meaning of a GLM solution) unanswered. By taking a different approach, based partly on Charnes, Cooper and Kortanek's semi-infinite programming duality (2), we shall resolve both of these difficulties. Another result we observe is that the Nemhauser-Widhelm conjecture on redundant cuts is false.

In addition, we investigate a related question of stability and demonstrate infimax value convergence even when multipliers (i.e., dual variables) do not exist.

PROGRAM DESCRIPTION

Let us now specify the problem under study and point out some basic properties. We suppose we are given the ordered set, (S, f, g, ρ), where

(1) S is a nonempty, compact set in R^n,

(2) f is a continuous functional on S,

(3) g is a continuous function from S into R^m,

(4) ρ is a relational "vector," where each ρ_i $(1 \leq i \leq m)$ is '$\leq$', '=', or '$\geq$'.

policies to be

$$S_w = \{x \text{ in } S : w(x) > 0\}.$$

Now we are ready to define the set of such functionals induced by S, namely,

$$W \equiv \{w: S \to [0,1] : |S_w| < \infty \text{ and } \sum_{x \in S_w} w(x) = 1\}.$$

Notice that the normalizing summation is well defined since there are only a finite number of used policies (i.e., $|S_w|$ is finite).

We next define a linear operator on the induced set of functionals which we call the "expected value" of a function, $F: S \to R^k$; that is,

$$E_w(F) = \sum_{x \in S_w} w(x) \ F(x).$$

The semi-infinite program induced by (S, f, g, ρ) is now defined to be

$$\text{Maximize } E_w(f) : w \text{ in } W \text{ and } E_w(g) \ \rho \ 0.$$

The decision elements in the above semi-infinite program are the functionals in W. One may interpret these as "mixed-strategy solutions" to an ordinary mathematical program of the form,

$$\text{Maximize } f(x) : x \text{ in } S \text{ and } g(x) \ \rho \ 0.$$

If we restrict the semi-infinite program to have $|S_w| = 1$, then we have a natural correspondance with the ordinary mathematical program, and such decisions are called "pure-strategy solutions." Thus, one may consider the semi-infinite program to be a relaxation (or linearization) of an ordinary mathematical program. It is easy to prove that GLM solves the semi-infinite program, and a gap prevails if, and only if, there is no pure strategy optimum.

When the ordinary program is convex, the well known saddle point theory applies which is equivalent to strong Lagrangian duality. Equivalently, one may note that every mixed-strategy, w, is dominated by the pure strategy having $x=E_w(X)$,

where X is the identity function, $X(x)=x$. When convexity structure is not present, weak Lagrangian duality still applies, and this is the foundation of GLM. The standard column generation method applied to the set of generated policies, say $\{x^k\}_k$, a set of associated columns is defined, and a sub-optimal mixed-strategy is found from ordinary linear programming. Using the associated dual variables the Lagrangian is maximized to generate x^{k+1} or deduce optimality. In the former case the process continues and may become of infinite duration.

Finite termination results in an optimal mixed-strategy solution and a minimal multiplier (associated with GLM). Our concern is with infinite duration and optimality of limit points. (It should be noted that tactical variations, such

For the case of all inequalities (where ρ_i is '$\leq$' for all i) the Nemhauser-Widhelm theorem says that if $g(x) < 0$ for some x in S, and if there is no duality gap, then convergence is assured. Their result (and proof) is essentially the same as Zangwill's (8) application of his cutting-plane theorem. Zangwill's assumption of convexity was only to set up the duality and dismiss the gap problem. His dual convergence proof does not rely upon any convexity structure in the primal. It should be noted that the Nemhauser-Widhelm theorem leaves dual convergence in doubt if a gap prevails. In the next section a more general theorem is presented to resolve these difficulties.

MAIN RESULTS

An examination of Zangwill's proof reveals that we need only be able to set up an appropriate dual where the column generation procedure (applied to the primal) corresponds to a cutting-plane procedure (applied to the dual). Then, it becomes

a matter of bounding the dual variables. (This is where Slater's condition is used.)

The dual of interest here is as follows:

$$\text{Minimize } z: \; y \text{ in } Y \text{ and}$$

$$z + y\,g(x) \geq f(x) \quad \text{for all } x \text{ in } S,$$

where $Y \equiv \{y \text{ in } R^m: 0\rho_i y_i \text{ for } i: \rho_i \text{ is not '='}\}$. Using the Charnes-Cooper-Kortanek theory (2) we need only establish "canonical closure" in order to deduce the strong duality relation. First, the compactness and continuity assumptions imply a compact range of (f,g) on S, so it only remains to construct a point in dual space to satisfy their "uniform Slater condition." This is achieved by setting $\bar{y} = 0$ and $\bar{z} = 1 + \text{Max}\{f(x); x \text{ in } S\}$, for then we have

$$\bar{z} + \bar{y}\,g(x) \geq f(x) + 1 \text{ for all } x \text{ in } S.$$

It now remains to provide a qualification on g that ensures a dual solution. To that end we define the

EXTENDED SLATER CONDITION:

There exists points in S for each $i=1,\ldots,m$ as follows:

(1) if ρ_i is '$\leq$', then there exists u^i such that $g(u^i)\rho 0$ and $g_i(u^i)<0$

(2) if ρ_i is '$\geq$', then there exists v^i such that $g(v^i)\rho 0$ and $g_i(v^i)>0$

(3) if ρ_i is '=', then there exists u^i and v^i such that

$$\left.\begin{array}{l} g_j(u^i)\rho 0 \\ g_j(v^i)\rho 0 \end{array}\right\} \text{ for } j \neq i$$

$$g_i(u^i) < 0 < g_i(v^i).$$

For all equality constraints only (3) applies, so the extended Slater Condition assumes the existance of 2m points with (3) holding. Since inequalities can be converted to equalities by adding a slack or surplus variable, there is no loss in generality to always assume 2m points satisfying (3). For example, if we have ρ_i as '$\leq$' and we convert to $g_i(x) + s^2 = 0$, then v^i exists by choosing s^i sufficiently large for an arbitrary x^i in S. Only existance of $u^i = (x^i, 0)$ remains to be demonstrated, where $g_i(x^i) < 0$.

It should be noted that for the case of all inequalities the Extended Slater Condition is weaker than the usual Slater condition since we may have the m points, $\{u^i\}_1^m$, and not have a single point, say x, such that $g(x)<0$. Of course, if g is explicitly quasi-convex, then the conditions are equivalent by defining $x = (\sum_{i=1}^{m} u^i) / m$.

We now prove that the Extended Slater Condition is sufficient to bound the multipliers without loss in optimality. First, observe that z can be bounded without loss in optimality (trivially), so there exists a constant, K, such that y may be restricted to

$$y\, g(x) \geq K \quad \text{for all } x \text{ in } S$$

without loss in optimality. In particular, by setting $x = u^i$ and then to v^i we obtain

$$K / g_i(v^i) \leq y_i \leq K / g_i(u^i).$$

Hence, Zangwill's theorem may now be applied to prove the

DUAL CONVERGENCE THEOREM:

Suppose $\{y^t\}_0^k$ are dual variables generated by the column generation procedure and g satisfies the Extended Slater Condition. Then,

(1) there exists at least one limit point of $\{y^t\}$,

(2) every limit point of $\{y^t\}$ minimizes $L^*(y) \equiv \text{Max}\{f(x) - y\,g(x) : x \text{ in } S\}$.

Thus, dual convergence is assured under mild assumptions (even when there is a Lagrangian duality gap). Notice that the Nemhauser-Widhelm example (5, p. 1055) cannot happen under these assumptions (which are implied by theirs). In fact, every cut generated in multiplier space by maximizing the Lagrangian supports the epigraph of L* and hence cannot be redundant as they conjectured. However, since the converse is also true (i.e., the cut is a dual support only if the Lagrangian is maximized), redundancy may arise from variation in pricing (i.e., not actually maximizing the Langrangian, but merely look for improvement).

Now let us consider convergence in the primal functional space. At each iteration we have an optimal mixed-strategy solution, w^k, to a relaxed problem which ignores those policies not generated or those which are dominated by the present set of generated policies. Since W is not compact, $\{w^k\}$ need not have a limit in W. However, let us suppose w^∞ is a limit point of $\{w^k\}$ and that w^∞ is in W. Further, suppose w* is an optimal mixed-strategy solution.

From elementary duality we have

$$E_{w^k}(f) = z^k \rightarrow \text{Inf}\{L^*(y)\} = E_{w^*}(f),$$

where w* is an optimal solution. Therefore, w^∞ is an optimal solution, and

PRIMAL CONVERGENCE THEOREM:

Under the assumptions of the Dual Convergence Theorem every limit point of $\{w^k\}$ is an optimal mixed strategy.

STABILITY

We now consider a relaxation of the Extended Slater Condition to require only feasibility and consider limits of perturbation. Namely, consider the perturbed program,

$$\text{Maximize } E_w(f): \quad E_w(g_i) \leq \varepsilon \text{ for } i:\rho_i \text{ is '}\leq\text{'}$$
$$E_w(g_i) \geq -\varepsilon \text{ for } i:\rho_i \text{ is '}\geq\text{'}$$
$$-\varepsilon \leq E_w(g_i) \leq \varepsilon \text{ for } i:\rho_i \text{ is '='}$$
$$\text{and } w \text{ in } W$$

where $\varepsilon > 0$. Our concern is with $\varepsilon \to 0_+$.

From the Dual Convergence theorem we know that a dual solution, say y^ε, exists (and may be approached) for any $\varepsilon \geq 0$.

First, consider the perturbed dual below:

$$\text{Minimize } z: \quad y \text{ in } Y \text{ and}$$
$$z + y(g(x) - \varepsilon d) = f(x) \quad \text{for all } x \text{ in } S,$$

$$\text{where } d_i = \begin{cases} 1 \text{ if } \rho_i \text{ is '}\leq\text{' or '='} \\ -1 \text{ if } \rho_i \text{ is '}\geq\text{'} \end{cases}$$

Define $q^\varepsilon = \varepsilon \sum_{i=1} y_i^\varepsilon$, so

$$z^\varepsilon = \text{Min}\left\{z: \; z + y\, g(x) \geq f(x) + q^\varepsilon \text{ for all } x \text{ in } S\right\}.$$

We have

$$\lim_{\varepsilon \to 0_+} z^\varepsilon = \lim_{\varepsilon \to 0+} \text{Inf}\left\{z: \; z + y\, g(x) \geq f(x) + q^\varepsilon \text{ for all } x \text{ in } S\right\}.$$

$$= \lim_{\varepsilon \to 0+} \left\{\text{Inf}\left\{L^*(y) + q^\varepsilon\right\}\right\}$$

$$= \text{Inf}\left\{L^*(y)\right\} + \lim_{\varepsilon \to 0+} q^\varepsilon$$

Therefore, $z^{\varepsilon} \to \text{Inf}\{L^*(y)\}$ if $q^{\varepsilon} \to 0$. To prove the latter recall the bounds derived earlier, so that there exists constants, K1 and K2, such that

$$|y_i^{\varepsilon}| = \frac{K1}{K2 - \varepsilon},$$

where $K2 \neq 0$ this implies

$$|q^{\varepsilon}| \leqq m \frac{K1\, \varepsilon}{K2 - \varepsilon}$$

so $q^{\varepsilon} \to 0$.

We have now proven the

PERTURBATION THEOREM:

Let y^{ε} solve the perturbed program. Then, for $\varepsilon \to 0_+$, the minimax value of the Lagrangian (i.e., $L^*(y^{\varepsilon})$) converges to the infimax value of the original program (i.e., Inf $L^*(y)$)

In closing it is also interesting to note the effect of approximation when maximizing the Lagrangian. Suppose we only generate ε-solutions so that x^k need only satisfy

$$f(x^k) - y\, g(x^k) \geqq L^*(y) - \varepsilon .$$

Then, this merely redefines the dual constraints to be

$$z + yg(x) \geqq f(x) - \varepsilon \text{ for all } x \text{ in } S.$$

It is easy to see that

$$z(\varepsilon) = \text{Inf}\{L^*(y) - \varepsilon\} = \text{Inf}\{L^*(y)\} - \varepsilon,$$

so $z(\varepsilon) = z(0) - \varepsilon$.

This means that the minimizing multipliers, if they exist, are still minimal in the dual approximation resulting from seeking ε-solutions when "maximizing" the Lagrangian.

REFERENCES

1. R. Brooks and A. Geoffrion, "Finding Everett's Lagrange Multipliers by Linear Programming," Opns. Res. 14, 1149-1153 (1966).

2. A. Charnes, W. W. Cooper and K. O. Kortanek, "On the Theory of Semi-Infinite Programming and a Generalization of the Kuhn-Tucker Saddle Point Theorem for Arbitrary Convex Functions," Nov. Res. Log. Quart. 16, 41-51 (1970).

3. H. Everett, "Generalized Lagrange Multiplier Method for Solving Problems of Optimum Allocation of Resources," Opns. Res. 11, 399-417 (1963).

4. H. J. Greenberg and W. P. Pierskalla, "Stability Theorems for Infinitely Constrained Mathematical Programs," to appear JOTA (1975).

5. F. H. Murphy, "Column Dropping Procedures for the Generalized Programming Algorithm," Mgt. Sci. 19, 1310-1321 (1973).

6. G. L. Nemhauser and W. B. Widhelm, "A Modified Linear Program for Columnar Methods in Mathematical Programming," Opns. Res. 19, 1051-1060 (1971).

7. M. Slater, "Lagrange Multipliers Revisited: A Contribution to Nonlinear Programming," Cowles Commission Discussion Paper Math. 403, 1950.

8. W. I. Zangwill, "Nonlinear Programming: A Unified Approach, Prentice-Hall, Englewood Cliffs, New Jersey, 1969